

SOME PECULIARITIES OF THE SPEECH FUNCTION IN A CHANGED GAS MEDIUM

V. S. Kuznetsov

In recent years the interest of researchers has increased in the study of the effect on the human and animal organism of a helium-oxygen mixture obtained by replacing atmospheric nitrogen by helium [1-3]. The substitution of helium for nitrogen has a very strong effect on the character of speech [2].

In the available literature a relatively small number of papers [5, 6], has been devoted to the problem of studying the function of speech during the respiration of special gas mixtures. In 1962 there appeared the paper of R. Beil [9], which was devoted to the study of the frequency spectrum of vowels of the English spoken language pronounced while breathing a helium-oxygen (HeO) mixture. Spectral analysis of the recorded speech showed that the frequency of the fundamental changed but little in the HeO medium, while the frequency of the first, second and third formants of all vowels increase at a ratio dependent on the composition of the gas medium.

In our experiments (I. Ya. Borshchevskiy and V. S. Kuznetsov), the peculiarities of speech formation while breathing an HeO mixture were studied in the course of two long experiments, of a total duration of 22 and 30 days, in which two test subjects participated. In contrast with the work of non-Soviet authors, the test subjects not only breathed, but were also continuously confined in the HeO atmosphere for 10 and 25 days. The principal measurements were made in the morning hours, every two days during the experiment; control experiments in a normal atmosphere were performed 1 and 3 days prior to the beginning of the experiment and after its conclusion.

During the experiment we studied the following characteristics of the speech function:

- 1) the features of the frequency composition of speech from the results of a spectral analyzer of the vowels of Russian spoken language a, o, i, y, u, e, both in the syllables da, do, di, my, ekh, and in the "standard phrase" paz, dva, tri;
- 2) the quality of speech formation, from articulation data, while reading lists of words;
- 3) the dynamics of the amplitude characteristics of speech.

The "standard phrase" used prior to the research is not generally accepted and does not reflect a statistical distribution of the sounds of Russian speech, but because of its simplicity it is widely used in special investigations.

/233

The "standard phrase" was pronounced by the test subjects in one breath 5-6 times, and varied sounds were kept standard for 1.5-3 sec. In pronouncing

the sounds the test subjects made use of three pickups: an MD-55 electromagnetic microphone having a comparatively even frequency characteristic; a DEMSh-1A noiseproof differential microphone, whose frequency characteristic has a rise of 6 db/octave from 300 to 1,000 Hz; and an LA-5 throat microphone. In order to determine the effect of aural self-monitoring on the quality of articulation in an HeO medium, in some cases the ears of the test subjects were covered by mufflers with earphones, through which white noise or music was transmitted. Some word lists were read by the test subjects at a whisper. The microphones and throat microphone were located together with the test subjects within the object in which the artificial atmosphere was created. The speech of the test subjects was recorded on a tape recorder having a smooth frequency characteristic of reproduction of ± 3 db.

The word lists were received by a team of trained auditors and the percentage of correctly received words was determined. Spectral analysis of the "standard phrase" and of the vowels was made by means of a spectreanalyzer with 1/3-octave filters; for this purpose a roll of electromagnetic tape with a record of the standard phrase or of the vowel being investigated was spliced together. The roll was played on a tape recorder; from the output of the recorder the voltage was transmitted to the input of the spectral analyzer. The automatic switching rate of the filters was chosen so as to transmit the signal from the roll twice into one filter. In analyzing the vowels we determined the magnitude of the peak voltage developed by the signal of each sound in the 1/3-octave filter.

In the two experiments a total of 95 lists of 25 words each were subjected to articulation studies. Of these, 61 lists were read while breathing the HeO mixture and 34 under ordinary conditions. Each of the 6 vowels used were pronounced by the test subjects 100 times, and of these, 75 times in the HeO medium.

EXPERIMENTAL RESULTS

As is well known, a speech wave is characterized by a specific spectral shape whose resonant peaks are known as formants. The most intense harmonic components of the acoustic oscillations are concentrated in the formant regions. The classical theory of speech formation ascribes to each voiced sound a particular place for its formants in the frequency scale, making it possible to identify the different vowels [6]. Spectral analysis of the sounds by means of the 1/3-octave filters cannot completely replace exact formant analysis, due to the comparatively large width of the filters. But the location of the formants can be estimated with good accuracy from the maxima in the spectral pattern. Given in the figure are examples of such a type of analysis of the vowels *a*, *o*, *u*, pronounced by the test subjects under ordinary conditions and in the HeO medium through the linear channel. In the spectrograms the intensity of the spectral components is plotted along the vertical in relative units (db); along the horizontal are the mean frequencies of the filters. From the figure it is evident that the first formant in the vowel *a* is located in the region of 600-800 Hz under ordinary speech-formation conditions, whereas it is shifted to the region of 800-1200 Hz in the HeO medium. The second formant is shifted from the 1000-Hz region to 2000 Hz, and the third from 2500 Hz to the region of 4000-5000 Hz. At the same time the intensity level of the third formant is considerably reduced (by 15 db). The frequency of the energy maximum of the vowel *o* increases by a

/234

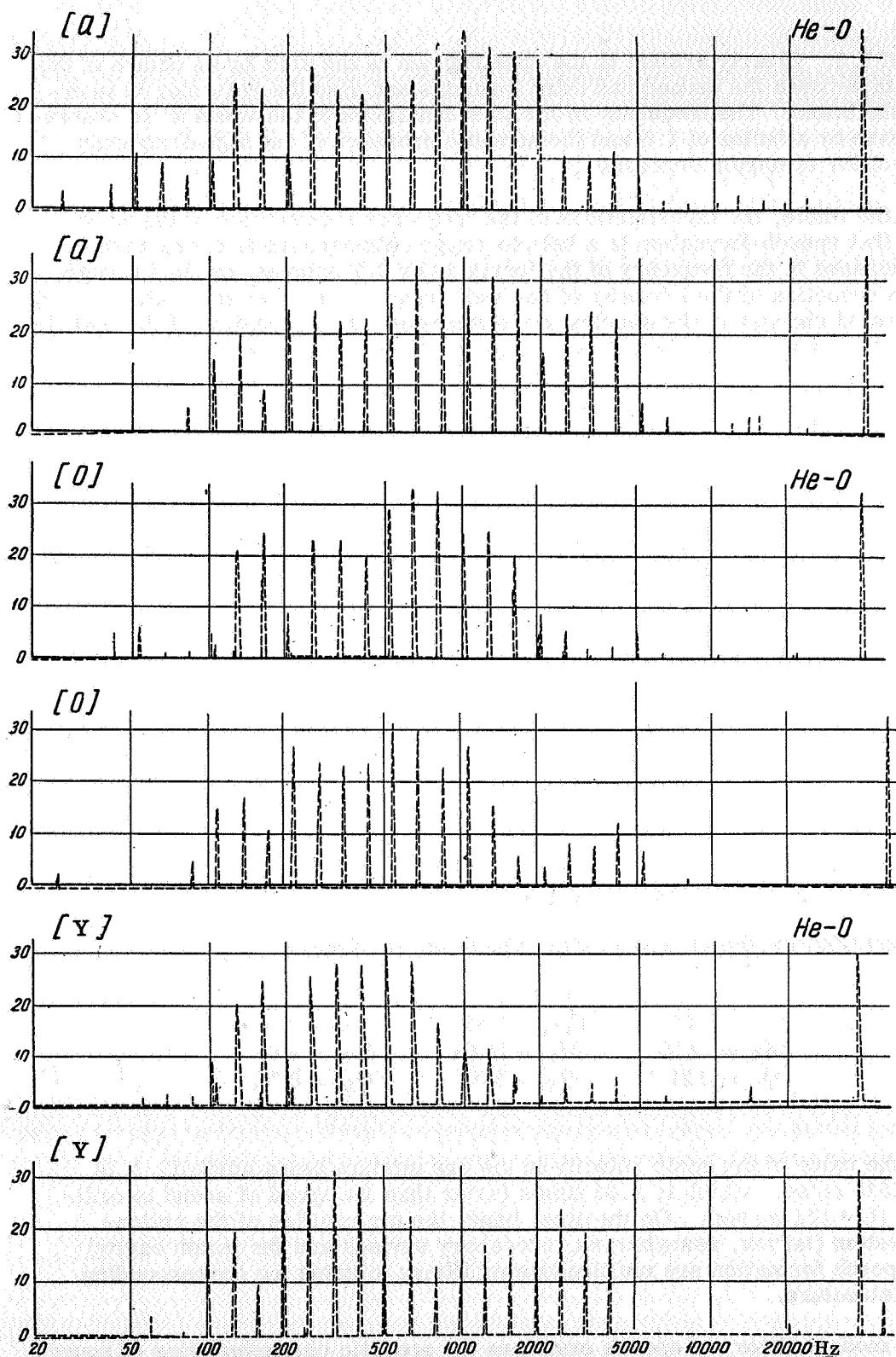


Figure 1. Example of Spectral Analysis of the Vowels a, o, y, Pronounced in a Helium-Oxygen (HeO) Atmosphere and in Normal Air.

factor of 1.3. Clearly evident in the spectrogram is the shift by an octave of the minimum between the second and third formants and also the reduction in intensity of the latter. The frequency of the first maximum of the vowel *u* is observed to increase by a factor of 1.5 and the acoustic pressure of the high-frequency portion of the spectrum decreases.

On the whole, the investigations of the frequency composition of the vowels showed that speech-formation in a helium-oxygen atmosphere is characterized by an increase in the frequency of the formants by 0.7 octaves, on the average, and by a reduction of the intensity of the high-frequency components. Subjectively, the indicated changes in the spectra are interpreted as a disruption of the melodic pattern of the voice, which takes on a "squeaky" character with shades of "nasality."

The increase of the formant frequencies of the vowels in the HeO atmosphere can be explained by the physical features of this gas mixture, in particular by the speed of propagation of the acoustic oscillations in the mixture and by the functional nature of the articulatory system. The artificial gas mixture used in the experiments contained 21% oxygen, 2-4% nitrogen and up to 75% helium. The rate of sound propagation in the gas mixture can be determined by the following formula:

$$C = \sqrt{\frac{RT^0 \left(\sum_i Q_i C_{Vi} + R \right)}{\sum_i M_i Q_i \sum_i Q_i C_{Vi}}},$$

where Q_i are the molar fractions of the different components (for ideal gases they coincide with the volumetric ratios); M_i , C_{Vi} are the molecular weights and the heat capacities of the respective components; R is a gas constant, equal to 1.987 cal/deg x mole or 8.314×10^7 erg x deg /mole [4].

By substituting the corresponding values into the formula:

/235

$C_{VO_2} = 4.96$	$C_{VN_2} = 4.94$	$C_{VHe} = 3.00$
$M_{O_2} = 32.00$	$M_{N_2} = 28.01$	$M_{He} = 4.00$
$Q_{O_2} = 0.21$	$Q_{N_2} = 0.04$	$Q_{He} = 0.75,$

we find the value of the sonic velocity in the gas mixture being studied. It is equal to 540 m/sec, which is 1.63 times faster than the speed of sound in ordinary air ($C = 331$ m/sec). On the other hand, the air cavities of the articulatory system (larynx, nasopharynx, accessory sinuses and the mouth cavity) during speech formation are multiresonator filters forming the corresponding formant structure.

The modern theory of speech examines the acoustic characteristics of sounds in the form of a product of the function (S), which represents modulation of the

air stream by the vocal chords, and of the filtering function of the vocal tract (T) [7]. The investigations of R. Beil show that the fundamental frequency of the voice, i. e., the oscillation frequency of the air stream does not change when breathing on HeO mixture. Consequently, the replacement of nitrogen by helium has no effect whatever on the functioning of the vocal chords. On the contrary, the function T, which is equivalent to several coupled resonators, may change considerably when the speed of propagation of sound increases. According to [9], the frequency of each formant is determined by the ratio of the sonic velocity to twice the length (l_n) of the corresponding equivalent resonator, i. e., $f_n =$

$C/2 l_n$. On the basis of this fact, it can be assumed that when the articulatory cavities are filled with an HeO mixture, the frequency of the formants increase, as compared with air, proportionally as the sonic velocity in the HeO medium exceeds the sonic velocity in air, i. e., by a factor of 1.63 in the case being considered, corresponding to a shift of 0.8 octaves toward the higher frequencies. Generalizing these conclusions, it can be said that speed formation in different gas mixtures is characterized by a shift in the location of the formants on the frequency scale, the shift being dependent on the ratio of the sonic velocity in the new medium to that in air. Frequency analysis of vowels in an HeO mixture has shown sufficiently good agreement of this conclusion with actual results. Without doubt, filling a resonator with the HeO mixture leads to a change of the entire transfer function of the speech tract; this, in particular, causes a reduction in the intensity of the high-frequency components of the spectrum. We did not make a special quantitative analysis of this phenomenon.

In studying the amplitude levels of speed it was found that a certain reduction in speech intensity (on the average 5 db) is observed immediately upon the transition to breathing the HeO mixture. On the second day of confinement of the test subject in the artificial atmosphere, the speech level increased almost to the original level owing to compensatory reactions and to aural self-control. Elimination of self-auditing by transmitting white noise and music to the earphones did not lead to a reduction in speech intensity or to a sharp deterioration of articulation (see the table). Sounds pronounced by the test subjects in the HeO medium are not only well identified, but also transmit certain personal traits of the speaker which are defined, as we know, by the fine structure of the spectrum. This is also confirmed by the articulatory changes. From the table it follows that the percentage of speech recognition when breathing the HeO mixture is close to 100. Especially fine are the results using the MD-55 and DEMS-1A microphones; somewhat worse is articulation using the laryngophone.

/236

It is necessary to emphasize, however, that the high percentage of recognition obtained in our experiments is explained by the good transmission conditions, by the high redundancy of the Russian language, and is supported by the considerable force of the aural and vocal compensation mechanisms. A proof of this may be the sharp reduction in speech recognition when reading word lists at a whisper (see the table). Elimination of the voice from the speech-forming process leads to a decrease in recognition from 95 to 50%. It must be assumed that in actual space flight the use of a helium-oxygen mixture may result in considerable deterioration of the communication quality in the presence of noise during transmission and reception if various speech correction measures are not provided for.

Recognizability of Speech (in %) From the Data of
Articulatory Measurements

Microphones	Air medium, days of test		Helium-oxygen medium, days of the test						
	1-st	2-nd	3-rd	5-th	9-th	18-th	24-th	29-th	
MD-55	100	100	100	100	95	95	100	100	
DEMSH-1A	95	100	100	100	100	100	100	100	
LA-5	85	85	70	70	90	80	55	80	
MD-55, whispering ..	95	95	50	50					
DEMSH-1A whispering	95	100	60	60					
MD-55 transmission . of white noise into the earphones	100	100	95	95	95	90			

REFERENCES

1. Boriskin, V. V. Gulyayev, P. A., Savin, B. M. In the collection Aviatsionnaya i kosmicheskaya meditsina [Aviation and Space Medicine]; Moscow, p. 78, 1963.
2. Dianov, A. G. Kosmicheskiye Issledovaniya; Vol. 2, p. 498, 1964.
3. Dianov, A. G. and Kuznetsov, A. G. In the collection Aviats. i. kosmich. med. [Aviat. and Space Med.]; Moscow, p. 162, 1963.
4. Kireyev, V. A. Kurs fizicheskoy khimii [Course in Physical Chemistry]; Phys. - Math. State Press, Moscow, 1951. /237
5. Mikhaylov, I. G., Solov'yev, V. A. and Skrnikov, Yu. P. Osnovy molekulyarnoy akustiki [Principles of Molecular Acoustics]; Nauka, Moscow, 1964.
6. Rzhavkin, S. N. Slukh i rech v svete sovremennykh fizicheskikh issledovaniy [Hearing and Speech in the Light of Modern Physiological Research]; Nauka, Publ. House, Moscow-Leningrad, 1936.
7. Fant, G. Akusticheskaya teoriya recheobrazovaniya [Acoustic Theory of Speech Formation]; Nauka, Moscow, 1964.
8. Barany, E. J. Acoust. Soc. Am. (JASA); Vol. 8, p. 217, 1937.
9. Beil, R. Ibid.; Vol. 34, p. 347, 1962.